Correlated Sources over Broadcast Channels

Paolo Minero and Young-Han Kim

Abstract—The problem of reliable transmission of correlated sources over the broadcast channel, originally studied by Han and Costa, is revisited. An alternative characterization of their sufficient condition for reliable transmission is given, which includes results of Marton for channel coding over broadcast channels and of Gray and Wyner for distributed source coding. A "minimalistic" coding scheme is presented, which is based on joint typicality encoding and decoding, without requiring the use of Cover's superposition coding, random hashing, and common part between two sources. The analysis of the coding scheme is also conceptually simple and relies on a new multivariate covering lemma and an application of the Fourier–Motzkin elimination procedure.

Fig. 1. The two-receiver broadcast channel with arbitrarily correlated sources.

I. INTRODUCTION

We study the problem of broadcasting two arbitrarily correlated sources over a general two-receiver discrete memoryless broadcast channel (DM-BC) as depicted in Figure 1. A discrete memoryless stationary source produces n independent copies $(S_1^n, S_2^n) \triangleq \{(S_{1,i}, S_{2,i})\}_{i=1}^n$ of a pair of random variables $(S_1, S_2) \sim p(s_1, s_2)$. The encoder maps the source sequences (S_1^n, S_2^n) into a sequence X^n from a finite set $\mathcal{X}$ and broadcasts it to two separate receivers. The communication channel is memoryless and modeled by a transition probability matrix $p(y_1, y_2|x)$, which maps each channel input x into a pair of finite valued output symbols (y_1, y_2). Decoder 1 maps the channel output sequence Y_1^n into an estimate $\hat{S}_1^n$ of the source sequence S_1^n. Similarly, decoder 2 maps Y_2^n into an estimate $\hat{S}_2^n$ of the source sequence S_2^n. For a given encoder and decoders, the average probability of error is defined as

$$\Pr\{(\hat{S}_1^n, \hat{S}_2^n) \neq (S_1^n, S_2^n)\}.$$

We say that the sources (S_1, S_2) can be reliably transmitted over the DM-BC $p(y_1, y_2|x)$ if there exists a sequence of encoders and decoder pairs such that the average probability of error vanishes as $n \to \infty$.

Han and Costa [16] provide a sufficient condition for reliable transmission of the sources (S_1, S_2) over a general DM-BC $p(y_1, y_2|x)$. Necessary conditions are derived by Gohari and Anantharam [14], and Kramer *et al.* [18]. Finding matching sufficient and necessary conditions is still an open problem in general.

Two special classes of sources and channels have been studied extensively in the literature. On the one hand, suppose $S_1 = (W_0, W_1)$ and $S_2 = (W_0, W_2)$, where W_0, W_1, and W_2 are three mutually independent random variables of entropies
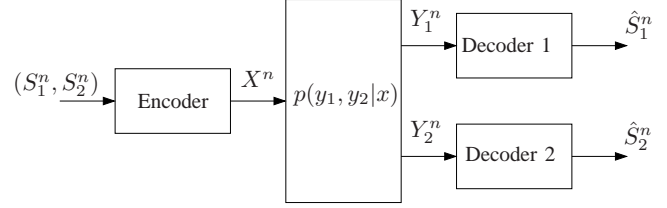
R_0, R_1, and R_2, respectively. The (closure of the) set of triples (R_0, R_1, R_2) which can be reliably transmitted is called the *capacity region* of the DM-BC. Inner bounds on the capacity region of a general DM-BC are developed in [3], [5], [13], [19], [21], [24], and the most general known achievable rate region is due to Marton [21], which is tight for several classes of broadcast channels. On the other hand, suppose the DM-BC is noiseless with $Y_1 = (X_0, X_1)$, $Y_2 = (X_0, X_2)$, where X_0, X_1, and X_2 have cardinalities 2^{R_0}, 2^{R_1}, and 2^{R_2}, respectively. The (closure of the) set of triples (R_0, R_1, R_2) which can be reliably transmitted is called the *distributed source coding region*, and its complete characterization is given by Gray and Wyner [15].

For transmission of arbitrarily correlated sources over a general noisy DM-BC, conclusive results are known for the case in which either source or channel has a degraded structure. More specifically, Han and Costa's sufficient condition matches known necessary conditions if decoder 1 is interested in recovering both source sequences (S_1^n, S_2^n) while decoder 2 is interested in S_2^n only (degraded source sets) [9] or if the DM-BC is degraded, or more generally, "more capable" [6], [18].

In this paper, we provide an alternative characterization to the sufficient condition of Han and Costa for reliable transmission of correlated sources. This new characterization, while algebraically equivalent to Han and Costa's original sufficient condition, is simpler and includes in a straightforward manner Marton's inner bound on the broadcast channel capacity region, Gray–Wyner distributed source coding region, and other aforementioned known results. In particular, the new characterization does not involve the common part of the two sources in the sense of Gács, Körner, and Witsenhausen [12], [26]. This implies that there is no special role played by the common part of the sources, confirming the standard engineering intuition. It is interesting to observe the difference to the problem of reliable transmission of correlated sources over *multiple access channels*. For this problem, the common part of the sources plays a pivotal role in Cover, El Gamal, and Salehi's random coding construction [4] by inducing coherent transmission from separate encoders.

The authors are with the Information Theory and Applications Center (ITA) of the California Institute of Telecommunications and Information Technologies (CALIT2), Department of Electrical and Computer Engineering, University of California, San Diego CA, 92093, USA. Email: pminero@ucsd.edu, yhk@ucsd.edu.

The highlight of the paper is a simple coding scheme for the sufficient condition and its analysis. This coding scheme is based on joint typicality encoding and decoding, and does not involve random hashing (Slepian–Wolf binning), Cover's superposition coding, or rate splitting as in the original coding scheme by Han and Costa. The performance of the coding scheme is analyzed via a new multivariate covering lemma and an application of the Fourier–Motzkin elimination procedure.

The rest of the paper is organized as follows. In the next section, we review Han and Costa's sufficient condition (with a recent correction by Kramer and Nair [11]). Section III presents the alternative characterization of the sufficient condition. The proof of this new coding theorem is given in Section IV. We conclude the paper in Section V with a discussion on implications of our coding scheme, and a sketch of yet another coding scheme for broadcasting correlated sources. Throughout the paper, we use the notation in [8].

II. HAN AND COSTA'S CODING THEOREM

For convenience, we recall here the coding theorem of Han and Costa with a few minor changes.

Theorem 1 ([16], [11]): The pair of sources (S_1, S_2) can be reliably transmitted over the DM-BC $p(y_1, y_2|x)$ if

$$H(S_1) < I(U_0, U_1, S_1; Y_1) - I(U_0, U_1; S_2|S_1),$$
$$H(S_2) < I(U_0, U_2, S_2; Y_2) - I(U_0, U_2; S_1|S_2),$$
$$\begin{aligned} H(S_1, S_2) < I(K, U_0, U_1, S_1; Y_1) + I(U_2, S_2; Y_2|K, U_0) \\ - I(U_1, S_1; U_2, S_2|K, U_0), \quad (1)\end{aligned}$$
$$\begin{aligned} H(S_1, S_2) < I(U_1, S_1; Y_1|K, U_0) + I(K, U_0, U_2, S_2; Y_2) \\ - I(U_1, S_1; U_2, S_2|K, U_0), \quad (2)\end{aligned}$$
$$\begin{aligned} H(S_1, S_2) < I(U_0, U_1, S_1; Y_1) + I(U_0, U_2, S_2; Y_2) \\ - I(U_1, S_1; U_2, S_2|K, U_0) - I(S_1, S_2; K, U_0), \\ (3)\end{aligned}$$

for some $p(u_0, u_1, u_2, x|s_1, s_2)$. Here $K = f(S_1) = g(S_2)$ denotes the common variable in the sense of Gács, Körner, and Witsenhausen, and the auxiliary random variable U_0 has the cardinality bound $|\mathcal{U}_0| \leq \min\{|\mathcal{X}||\mathcal{S}_1||\mathcal{S}_2| + 4, |\mathcal{Y}_1||\mathcal{Y}_2||\mathcal{S}_1||\mathcal{S}_2| + 4\}$.

The coding scheme by Han and Costa requires three auxiliary random variables (U_0, U_1, U_2). Roughly speaking, U_0 serves as a "cloud center" distinguishable by both receivers and carries the common part K and hash indexes of (S_1, S_2), U_1 and U_2 are codewords within the cloud centers which encode the remaining uncertainty about S_1 and S_2, respectively. Finally, Marton's subcode generation technique [21] is employed to obtain arbitrary correlation among (U_0, U_1, U_2). For details of this interpretation, see [16].

III. AN ALTERNATIVE CHARACTERIZATION

We have the following:

Theorem 2: The sources (S_1, S_2) can be reliably transmitted over the DM-BC $p(y_1, y_2|x)$ if

$$H(S_1) < I(U_0, U_1, S_1; Y_1) - I(U_0, U_1; S_2|S_1),$$
$$H(S_2) < I(U_0, U_2, S_2; Y_2) - I(U_0, U_2; S_1|S_2),$$

$$\begin{aligned} H(S_1, S_2) < I(U_0, U_1, S_1; Y_1) + I(U_2, S_2; Y_2|U_0) \\ - I(U_1, S_1; U_2, S_2|U_0), \quad (4)\end{aligned}$$
$$\begin{aligned} H(S_1, S_2) < I(U_1, S_1; Y_1|U_0) + I(U_0, U_2, S_2; Y_2) \\ - I(U_1, S_1; U_2, S_2|U_0), \quad (5)\end{aligned}$$
$$\begin{aligned} H(S_1, S_2) < I(U_0, U_1, S_1; Y_1) + I(U_0, U_2, S_2; Y_2) \\ - I(U_1, S_1; U_2, S_2|U_0) - I(S_1, S_2; U_0). \quad (6)\end{aligned}$$

for some $p(u_0, u_1, u_2, x|s_1, s_2)$. Here the auxiliary random variable U_0 has the cardinality bound $|\mathcal{U}_0| \leq \min\{|\mathcal{X}||\mathcal{S}_1||\mathcal{S}_2| + 4, |\mathcal{Y}_1||\mathcal{Y}_2||\mathcal{S}_1||\mathcal{S}_2| + 4\}$.

Remark 1: It is easy to see that Theorem 1 is a special case of Theorem 2. In fact, if we define a new random variable as $\tilde{U}_0 = (U_0, K)$, where K denotes the common part of the sources (S_1, S_2), then the inequalities in Theorem 2 for the triple $(\tilde{U}_0, U_1, U_2)$ simplify and reduce to those in Theorem 1. Conversely, the following proposition proves that Theorem 2 is a special case of Theorem 1, establishing the equivalence of the two.

Proposition 3: The sufficient condition for reliable transmission in Theorem 1 is equivalent to the one in Theorem 2.

Proof: See Appendix A. ∎

Remark 2: No cardinality bounds are known for the auxiliary random variables U_1 and U_2.

Remark 3: Application of Theorem 2 yields the following results as special cases:

a) *Marton's inner bound [21]:* Consider the special case of independent sources described above, so take $S_1 = (W_0, W_1)$ and $S_2 = (W_0, W_2)$, with W_0, W_1, and W_2 of entropies R_0, R_1, R_2, respectively. By choosing (U_0, U_1, U_2) to be independent of (W_0, W_1, W_2), Theorem 2 yields Marton's inner bound, which state that the capacity region of the DM-BC $p(y_1, y_2|x)$ contains the set of rate triples (R_0, R_1, R_2) satisfying

$$R_0 + R_1 < I(U_0, U_1; Y_1),$$
$$R_0 + R_2 < I(U_0, U_2; Y_2),$$
$$\begin{aligned} R_0 + R_1 + R_2 < I(U_0, U_1; Y_1) + I(U_2; Y_2|U_0) \\ - I(U_1; U_2|U_0),\end{aligned}$$
$$\begin{aligned} R_0 + R_1 + R_2 < I(U_1; Y_1|U_0) + I(U_0, U_2; Y_2) \\ - I(U_1; U_2|U_0),\end{aligned}$$
$$\begin{aligned} 2R_0 + R_1 + R_2 < I(U_0, U_1; Y_1) + I(U_0, U_2; Y_2) \\ - I(U_1; U_2|U_0)\end{aligned}$$

for some $p(u_0, u_1, u_2, x)$.

b) *The Gray–Wyner source coding problem [15]:* Consider the noiseless DM-BC channel with $Y_1 = (X_0, X_1)$, $Y_2 = (X_0, X_2)$ with links of rate $R_0 = H(X_0)$, $R_1 = H(X_1)$, $R_2 = H(X_2)$. By taking $U_1 = X_1$, $U_2 = X_2$, and $U_0 = (X_0, V)$, under the distribution $p(v|s_1, s_2)p(x_0)p(x_1)p(x_2)$, the inequalities in Theorem 2 simplify to the following rate region

$$R_0 + R_1 > I(S_1, S_2; V) + H(S_1|V),$$
$$R_0 + R_2 > I(S_1, S_2; V) + H(S_2|V),$$
$$R_0 + R_1 + R_2 > I(S_1, S_2; V) + H(S_1|V) + H(S_2|V),$$
$$2R_0 + R_1 + R_2 > 2I(S_1, S_2; V) + H(S_1|V)$$

$$+ H(S_2|V)$$

which includes (and is in fact equivalent to) the Gray–Wyner distributed source coding region, which is characterized by the following set of inequalities:

$$R_0 > I(S_1, S_2; V),$$
$$R_1 > H(S_1|V),$$
$$R_2 > H(S_2|V).$$

c) *Degraded "source" sets*: Consider the case in which decoder 1 is interested in reconstructing both source sequences (S_1^n, S_2^n) with vanishing error probability. This setup can be captured by considering transmission of a pair of sources $\tilde{S}_1 = (S_1, S_2)$, $\tilde{S}_2 = S_2$. By choosing $U_0 = (U, S_2)$, $U_1 = X$, and $U_2 = $ const. with $p(u, x|s_1, s_2) = p(u, x)$, the conditions stated in the Theorem 2 reduce to

$$H(S_2) \leq I(U; Y_2),$$
$$H(S_1, S_2) \leq I(U; Y_2) + I(X; Y_1|U), \qquad (7)$$
$$H(S_1, S_2) \leq I(X; Y_1).$$

Conversely, [18] shows that if the sources (S_1, S_2) can be reliably transmitted in this setup, then conditions (7) with "<" replaced by "≤" have to be simultaneously satisfied for some $p(u, x)$. Thus, we conclude that the sufficient condition in Theorem 2 is also necessary for this class of problems. It should be remarked that the same condition can also be obtained by separately performing source and channel coding, first compressing the sources and then encoding the resulting sequences using the channel coding scheme in [10].

d) *More capable broadcast channels*: A DM-BC is said to be more capable if $I(X; Y_1) \geq I(X; Y_2)$ for all $p(x)$ (cf. [6]). If we apply Theorem 2 to this setting, and we choose $U_0 = (U, S_2)$, $U_1 = X$, and $U_2 = $ const., with $p(u, x|s_1, s_2) = p(u, x)$, then the conditions stated in the theorem reduce to (7). Conversely, it is shown by Kang and Kramer [9] that if the sources (S_1, S_2) can be reliably transmitted in this setup, then conditions (7) with "<" replaced by "≤" have to be simultaneously satisfied for some $p(u, x)$. Hence, the sufficient condition in Theorem 2 is also necessary for this class of problems, but again the same condition can be obtained by performing source and channel coding separately.

IV. PROOF OF THEOREM 2

The coding technique used to prove Theorem 2 can be outlined as follows. The scheme requires three auxiliary random variables (U_0, U_1, U_2). Information about the source sequence S_1 is carried by U_0 and U_1, while information about the source sequence S_2 is carried by U_0 and U_2. Although U_0 carries information about both source sequences, it is not treated as cloud center in our coding scheme, so an error in decoding U_0 does not preclude successful decoding of S_1 or S_2. Marton's subcode generation technique is used to obtain arbitrary correlation among (U_0, U_1, U_2).

The following definitions are needed for the remainder of the paper.

Definition 4: Let (x^n, y^n) be a pair of sequences with elements drawn from a pair of finite alphabets $(\mathcal{X}, \mathcal{Y})$. Define their joint empirical probability mass function as

$$\pi(x, y|x^n, y^n) \triangleq \frac{|\{i : (x_i, y_i) = (x, y)\}|}{n}, \text{ for } (x, y) \in \mathcal{X} \times \mathcal{Y}$$

Then, the set $\mathcal{T}_\epsilon^{(n)}(X, Y)$ (in short, $\mathcal{T}_\epsilon^{(n)}$) of jointly ϵ-typical n-sequences (x^n, y^n) is defined as:

$$\mathcal{T}_\epsilon^{(n)}(X, Y) \triangleq \{(x^n, y^n) : |\pi(x, y|x^n, y^n) - p(x, y)| \leq \epsilon p(x, y)$$
$$\text{for all } (x, y) \in \mathcal{X} \times \mathcal{Y}\}.$$

Similarly, the set $\mathcal{T}_\epsilon^{(n)}(Y|x^n)$ of ϵ-typical n-sequences y^n that are jointly typical with a given sequence x^n is defined as:

$$\mathcal{T}_\epsilon^{(n)}(Y|x^n) \triangleq \{y^n : (x^n, y^n) \in \mathcal{T}_\epsilon^{(n)}(X, Y)\}.$$

In the remaining of this section, we first describe the random codebook generation and encoding-decoding scheme, then we outline the analysis of the probability of error, which is treated in detail in Appendix C.

Random codebook generation: Let $\epsilon' > 0$. Fix a joint distribution $p(u_0, u_1, u_2|s_1, s_2)$ and, without loss of generality, let $p(x|u_0, u_1, u_2, s_1, s_2)$ be a chosen deterministic function $x(s_1, s_2, u_0, u_1, u_2)$. Compute $p(u_0)$, $p(u_1|s_1)$, and $p(u_2|, s_2)$ for the given source distribution $p(s_1, s_2)$. Randomly and independently generate 2^{nR_0} sequences $u_0^n(m_0)$, $m_0 \in [1 : 2^{nR_0}]$, each according to $\prod_{i=1}^n p_{U_0}(u_{0i})$. For each source sequence s_1^n randomly and independently generate 2^{nR_1} sequences $u_1^n(s_1^n, m_1)$, $m_1 \in [1 : 2^{nR_1}]$, according to $\prod_{i=1}^n p_{U_1|S_1}(u_{1i}|s_{1i})$. The same procedure, using $\prod_{i=1}^n p_{U_2|S_2}(u_{2i}|s_{2i})$, is repeated for generating 2^{nR_2} sequences $u_2^n(s_2^n, m_2)$, $m_2 \in [1 : 2^{nR_2}]$. The rates (R_0, R_1, R_2) are chosen so that the ensemble of generated sequences (u_0^n, u_1^n, u_2^n) "cover" the set $\mathcal{T}_{\epsilon'}^{(n)}(U_0, U_1, U_2|s_1^n, s_2^n)$ for all $(s_1^n, s_2^n) \in \mathcal{T}_{\epsilon'}^{(n)}(S_1, S_2)$. The conditions for "covering" are given in the following Lemma.

Lemma 5: Define the event

$$\mathcal{A} = \{(S_1^n, S_2^n, U_0^n(m_0), U_1^n(S_1^n, m_1),$$
$$U_2^n(S_2^n, m_2) \notin \mathcal{T}_{\epsilon'}^{(n)} \text{ for all } m_0, m_1, m_2\}.$$

Then, there exists $\delta(\epsilon')$ which tends to zero as $\epsilon' \to 0$ such that $\mathsf{P}(\mathcal{A}) \to 0$ as $n \to \infty$ if the following inequalities are satisfied

$$R_0 > I(U_0; S_1, S_2) + \delta(\epsilon'),$$
$$R_1 > I(U_1; S_2|S_1) + \delta(\epsilon'),$$
$$R_2 > I(U_2; S_1|S_2) + \delta(\epsilon'),$$
$$R_0 + R_1 > I(U_0; S_1, S_2) + I(U_1; U_0, S_2|S_1) + \delta(\epsilon'),$$
$$R_0 + R_2 > I(U_0; S_1, S_2) + I(U_2; U_0, S_1|S_2) + \delta(\epsilon'),$$
$$R_1 + R_2 > I(U_1, S_1; U_2, S_2) - I(S_1; S_2) + \delta(\epsilon'),$$
$$R_0 + R_1 + R_2 > I(U_0; S_1, S_2) + I(U_1; U_2|U_0, S_1, S_2)$$
$$+ I(U_1; U_0, S_2|S_1)$$
$$+ I(U_2; U_0, S_1|S_2) + \delta(\epsilon'). \qquad (8)$$

Proof: See Appendix B. ∎

Encoding: For each $(s_1^n, s_2^n) \in \mathcal{T}_\epsilon^{(n)}(S_1, S_2)$, choose a triple $(m_0, m_1, m_2) \in [1 : 2^{nR_0}] \times [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$

such that $(s_1^n, s_2^n, u_0^n(m_0), u_1^n(s_1^n, m_1), u_2^n(s_2^n, m_2)) \in \mathcal{T}_{\epsilon'}^{(n)}(S_1, S_2, U_0, U_1, U_2)$. Choose $(m_0, m_1, m_2) = (1, 1, 1)$ if no such triple can be found. Then, at time $i \in \{1, \ldots, n\}$, the encoder transmits $x_i = x(s_{1i}, s_{2i}, u_{0i}(m_0), u_{1i}(s_1^n, m_1), u_{2i}(s_2^n, m_2))$. The sequence x^n so generated is the codeword corresponding to the source sequence (s_1^n, s_2^n).

Decoding: Let $\epsilon > \epsilon'$. Upon observing the sequence y_1^n, decoder 1 declares that s_1^n is the transmitted source sequence if this is the unique sequence such that $(s_1^n, u_0^n(m_0), u_1^n(s_1^n, m_1), y_1^n) \in \mathcal{T}_{\epsilon}^{(n)}(S_1, U_0, U_1, Y_1)$ for some $m_0 \in [1 : 2^{nR_0}]$ and $m_1 \in [1 : 2^{nR_1}]$. Otherwise, declare an error. Similarly, the decoder observing y_2^n declares that s_2^n was sent, if this is the unique sequence such that $(s_2^n, u_0^n(m_0), u_2^n(s_2^n, m_2), y_2^n) \in \mathcal{T}_{\epsilon}^{(n)}(S_2, U_0, U_2, Y_2)$ for some $m_0 \in [1 : 2^{nR_0}]$ and $m_2 \in [1 : 2^{nR_2}]$.

Error events: The detailed analysis of the error probability is given in Appendix C, but it can be outlined as follows. If the "covering" conditions in Lemma 5 are satisfied, then the probability of error encoding can be made arbitrarily small by letting $n \to \infty$. Thus, we focus on the analysis of decoding errors. Assume that $(m_0, m_1, m_2) = (M_0, M_1, M_2)$ is selected for the source sequence (s_1^n, s_2^n). The error event for decoder 1 can be divided into two parts:

1) There exists a sequence $\hat{s}_1^n \neq S_1^n$ such that $(\hat{s}_1^n, u_0^n(M_0), u_1^n(\hat{s}_1^n, m_1), y_1^n) \in \mathcal{T}_{\epsilon}^{(n)}(S_1, U_0, U_1, Y_1)$ for some $m_1 \in [1 : 2^{nR_1}]$. The probability of this event vanishes as $n \to \infty$ if the following inequality is satisfied

$$H(S_1) + R_1 < I(U_1, S_1; U_0, Y_1) - 2\delta(\epsilon). \quad (9)$$

2) There exists a sequence $\hat{s}_1^n \neq S_1^n$ and an $m_0 \neq M_0$ such that $(\hat{s}_1^n, u_0^n(m_0), u_1^n(\hat{s}_1^n, m_1), y_1^n) \in \mathcal{T}_{\epsilon}^{(n)}(S_1, U_0, U_1, Y_1)$ for some $m_1 \in [1 : 2^{nR_1}]$. The probability of this event can be made arbitrarily small as $n \to \infty$ if

$$H(S_1) + R_0 + R_1 < I(U_0, U_1, S_1; Y_1) \\ + I(U_0; U_1, S_1) - 3\delta(\epsilon). \quad (10)$$

Similarly, the probability of error for decoder 2 can be made arbitrarily small as n tends to infinity if

$$H(S_2) + R_2 < I(U_2, S_2; U_0, Y_2) - 2\delta(\epsilon), \quad (11)$$

$$H(S_2) + R_0 + R_2 < I(U_0, U_2, S_2; Y_2) \\ + I(U_0; U_2, S_2) - 3\delta(\epsilon). \quad (12)$$

Fourier–Motzkin elimination: In summary, by combining the inequalities in Lemma 5, (9), (10), (11), and (12), and by letting $\epsilon', \epsilon \to 0$, we conclude that the sources (S_1, S_2) can be reliably transmitted over the DM-BC $p(y_1, y_2 | x)$ if there exists a rate tuple (R_0, R_1, R_2) satisfying

$$R_0 > I(U_0; S_1, S_2) \triangleq v_1,$$
$$R_1 > I(U_1; S_2 | S_1) \triangleq v_2,$$
$$R_2 > I(U_2; S_1 | S_2) \triangleq v_3,$$
$$R_0 + R_1 > I(U_0; S_1, S_2) + I(U_1; U_0, S_2 | S_1) \triangleq v_4,$$
$$R_0 + R_2 > I(U_0; S_1, S_2) + I(U_2; U_0, S_1 | S_2) \triangleq v_5,$$

$$R_1 + R_2 > I(U_1, S_1; U_2, S_2) - I(S_1; S_2) \triangleq v_6,$$
$$R_0 + R_1 + R_2 > I(U_0; S_1, S_2) + I(U_1; U_2 | U_0, S_1, S_2) \\ + I(U_1; U_0, S_2 | S_1) \\ + I(U_2; U_0, S_1 | S_2) \triangleq v_7.$$
$$H(S_1) + R_1 < I(U_1, S_1; U_0, Y_1) \triangleq v_8,$$
$$H(S_1) + R_0 + R_1 < I(U_0, U_1, S_1; Y_1) + I(U_0; U_1, S_1) \triangleq v_9,$$
$$H(S_2) + R_2 < I(U_2, S_2; U_0, Y_2) \triangleq v_{10},$$
$$H(S_2) + R_0 + R_2 < I(U_0, U_2, S_2; Y_2) + I(U_0; U_2, S_2) \triangleq v_{11}$$
$$(13)$$

for some $p(u_0, u_1, u_2, x | s_1, s_2)$. The final step in our derivation consists of eliminating the auxiliary variables R_0, R_1, R_2 from the above set of inequalities. This routine task can be performed by application of the standard Fourier–Motzkin elimination algorithm [17] [27]. Unfortunately, brute force application of this algorithm to eliminate R_0, R_1, R_2 (in this order) results in 28 inequalities involving $H(S_1)$, $H(S_2)$, and $H(S_1, S_2)$, most of which are redundant because implied by other inequalities. Also, to find a minimal set of non-redundant inequalities, one has to verify linear information inequalities involving several random variables. Therefore, although conceptually simple, application of the Fourier–Motzkin elimination algorithm can be tedious.

We would like to illustrate a technique that allow us to get around these difficulties and to efficiently perform the elimination using a computer program. The key idea is to treat the right hand sides of (13) as auxiliary variables. Denote them by $v_1, \ldots, v_{11}$. Next, observe that

$$\begin{aligned} v_1 &\geq 0, \\ v_2 &\geq 0, \\ v_3 &\geq 0, \\ v_1 + v_2 &\leq v_4, \\ v_1 + v_3 &\leq v_5, \\ v_2 + v_3 &\leq v_6, \\ v_4 + v_6 &\leq v_2 + v_7, \\ v_4 + v_5 &\leq v_1 + v_7, \\ v_5 + v_6 &\leq v_3 + v_7, \\ v_8 &\leq v_9, \\ v_{10} &\leq v_{11}. \end{aligned} \quad (14)$$

The above information inequalities can be checked directly using the chain rule and the nonnegativity of mutual information or, alternatively, they can be verified by the software ITIP [25].

Combining (13) and (14) we obtain a system of linear inequalities involving the variables $(H(S_1), H(S_2), R_0, R_1, R_2, v_1, v_2, \ldots, v_{11})$. Next, we eliminate the auxiliary variables R_0, R_1 and R_2 from this system of linear inequalities. Since (13) and (14) have constant coefficients, the Fourier–Motzkin elimination can now be performed by a computer program, e.g., by the software PORTA [1]. The algorithm results in the following inequalities involving $H(S_1)$ and $H(S_2)$:

$$H(S_1) < v_9 - v_4,$$
$$H(S_1) < v_8 - v_2,$$
$$H(S_2) < v_{11} - v_5,$$

$$H(S_2) < v_{10} - v_3,$$
$$H(S_1) + H(S_2) < v_9 + v_{10} - v_7,$$
$$H(S_1) + H(S_2) < v_8 + v_{11} - v_7,$$
$$H(S_1) + H(S_2) < v_8 + v_{10} - v_6,$$
$$H(S_1) + H(S_2) < v_9 + v_{11} - v_7 - v_1. \tag{15}$$

To complete the proof, we argue that three of the above inequalities can be discarded because inactive. We proceed as follows. Define three new auxiliary random variables $\tilde{U}_i = (U_i, W)$, $i \in \{0, 1, 2\}$, where W is chosen independent of everything else, and replace U_i with $\tilde{U}_i$ in (13). Substituting the new values of $v_1, \ldots, v_{11}$, and performing some manipulations, we can rewrite (15) as follows

$$H(S_1) < I(U_0, U_1, S_1; Y_1) - I(U_0, U_1; S_2|S_1),$$
$$H(S_1) < I(U_1, S_1; U_0, Y_1)$$
$$\qquad - I(U_1; S_2|S_1) + H(W), \tag{16}$$
$$H(S_2) < I(U_0, U_2, S_2; Y_2) - I(U_0, U_2; S_1|S_2),$$
$$H(S_2) < I(U_2, S_2; U_0, Y_2)$$
$$\qquad - I(U_2; S_1|S_2) + H(W), \tag{17}$$
$$H(S_1, S_2) < I(U_0, U_1, S_1; Y_1) + I(U_2, S_2; Y_2|U_0),$$
$$\qquad - I(U_1, S_1; U_2, S_2|U_0),$$
$$H(S_1, S_2) < I(U_1, S_1; Y_1|U_0) + I(U_0, U_2, S_2; Y_2)$$
$$\qquad - I(U_1, S_1; U_2, S_2|U_0),$$
$$H(S_1, S_2) < I(U_2, S_2; U_0, Y_2) + I(U_1, S_1; U_0, Y_1)$$
$$\qquad - I(U_1, S_1; U_2, S_2) + H(W), \tag{18}$$
$$H(S_1, S_2) < I(U_0, U_1, S_1; Y_1) + I(U_0, U_2, S_2; Y_2)$$
$$\qquad - I(U_1, S_1; U_2, S_2|U_0) - I(S_1, S_2; U_0).$$

By letting $H(W) \to \infty$, we can make (16), (17) and (18) trivial, and thus inactive, while leaving the remaining inequalities unchanged. In summary, the five inequalities remaining from the Fourier–Motzkin elimination are those appearing in the statement of Theorem 2. This proves the desired sufficiency.

V. DISCUSSION

The best known inner bound for the problem of broadcasting correlated sources over a general DM-BC is due to Han and Costa with the associated coding scheme that cleverly combines a mélange of coding techniques, including joint typicality encoding and decoding, random hashing, superposition coding, and the use of the common part between two random variables. To investigate which techniques are crucial, this paper presents a "minimalistic" coding scheme in which we remove from the coding scheme of Han and Costa all unnecessary components without affecting the overall performance. Our proposed coding scheme does not require random hashing and superposition coding, and it does not involve the common part of two random variables. This highlights how source encoding can be performed by simply jointly "covering" the set of typical sources using auxiliary correlated random variables. An interesting implication of our result, albeit no more than theoretically amusing, is that the capacity of the degraded DM-BC can be achieved without employing Cover's superposition coding [2].

We would like to conclude the paper by showing that the sufficient condition stated in Theorem 2 can also be proved by means of a more elaborate coding scheme which uses superposition coding in addition to joint typicality encoding and decoding. This alternative scheme differs from the one described in Section IV in the way the random codebook is generated, as the auxiliary random variable U_0 serves as a cloud center for generating the auxiliary random variables U_1 and U_2. Appendix D presents a description of the code construction and a sketch of the analysis of the associated probability of error.

APPENDIX A
PROOF OF PROPOSITION 3

Suppose that the source (S_1, S_2) satisfies the sufficient condition stated in Theorem 2, so that the inequalities in Theorem 2 are satisfied for some triple (U_0, U_1, U_2). The goal is to show that (S_1, S_2) also satisfy the inequalities in Theorem 1 for the same triple of auxiliary random variables. Expanding the left hand side of (4), we obtain that

$$H(S_1, S_2)$$
$$< I(U_0, U_1, S_1; Y_1) + I(U_2, S_2; Y_2|U_0)$$
$$\quad - I(U_1, S_1; U_2, S_2|U_0),$$
$$= H(Y_1) - H(Y_1|U_0, U_1, S_1) - H(S_2, U_2|Y_2, U_0)$$
$$\quad + H(S_2, U_2|U_0, U_1, S_1)$$
$$\leq H(Y_1) - H(Y_1|K, U_0, U_1, S_1) - H(S_2, U_2|Y_2, K, U_0)$$
$$\quad + H(S_2, U_2|K, U_0, U_1, S_1),$$
$$= I(K, U_0, U_1, S_1; Y_1) + I(U_2, S_2; Y_2|K, U_0)$$
$$\quad - I(U_1, S_1; U_2, S_2|K, U_0),$$

where the second inequality follows from the fact that conditioning reduces the entropy and that $K = f(S_1) = g(S_2)$ is a deterministic function of the sources (S_1, S_2). Thus, (S_1, S_2) satisfy (1). Proceeding in a similar way, it is immediate to show that (S_1, S_2) satisfy (2). Finally, we expand the left hand side of (6) and obtain that

$$H(S_1, S_2)$$
$$< I(U_0, U_1, S_1; Y_1) + I(U_0, U_2, S_2; Y_2)$$
$$\quad - I(U_1, S_1; U_2, S_2|U_0) - I(S_1, S_2; U_0).$$
$$= H(Y_1) - H(Y_1|U_0, U_1, S_1) - H(Y_2) - H(Y_2|U_0, U_2, S_2)$$
$$\quad - H(U_1|S_1, U_0) + H(U_1, S_1|U_0, U_2, S_2) - H(S_1)$$
$$\leq H(Y_1) - H(Y_1|K, U_0, U_1, S_1) - H(Y_2)$$
$$\quad - H(Y_2|K, U_0, U_2, S_2) - H(U_1|S_1, K, U_0)$$
$$\quad + H(U_1, S_1|K, U_0, U_2, S_2) - H(S_1)$$
$$= I(K, U_0, U_1, S_1; Y_1) + I(K, U_0, U_2, S_2; Y_2)$$
$$\quad - I(U_1, S_1; U_2, S_2|K, U_0) - I(S_1, S_2; K, U_0),$$

so the sources (S_1, S_2) satisfy (3). Thus, we conclude that (S_1, S_2) also satisfy the sufficient condition stated in Theorem 1.

APPENDIX B
PROOF OF LEMMA 5

The technique used in the proof is similar to the one in [7]. For each $(s_1^n, s_2^n) \in \mathcal{T}_{\epsilon'}^{(n)}(S_1, S_2)$, let

$$\mathcal{A}(s_1^n, s_2^n)$$
$$\triangleq \{(m_0, m_1, m_2) \in [1:2^{nR_0}] \times [1:2^{nR_1}] \times [1:2^{nR_2}] :$$
$$(s_1^n, s_2^n, U_0^n(m_0), U_1^n(s_1^n, m_1), U_2^n(s_2^n, m_2)) \in \mathcal{T}_{\epsilon'}^{(n)}\}.$$

Then,

$$P(\mathcal{A}) \leq P((s_1^n, s_2^n) \notin \mathcal{T}_{\epsilon'}^{(n)})$$
$$+ \sum_{(s_1^n, s_2^n) \in \mathcal{T}_{\epsilon'}^{(n)}} p(s_1^n, s_2^n) P(|\mathcal{A}(s_1^n, s_2^n)| = 0).$$

By the law of large number the first term in the sum tends to 0 as $n \to \infty$. To bound the second term, observe that

$$P(|\mathcal{A}(s_1^n, s_2^n)| = 0) \leq \frac{\text{Var}(|\mathcal{A}(s_1^n, s_2^n)|)}{(E[|\mathcal{A}(s_1^n, s_2^n)|])^2}. \qquad (19)$$

Using indicator random variables, we express $|\mathcal{A}(s_1^n, s_2^n)|$ as

$$|\mathcal{A}(s_1^n, s_2^n)| = \sum_{m_0, m_1, m_2} \mathbb{1}_{\mathcal{A}_{m_0, m_1, m_2}}$$

where

$$\mathcal{A}_{m_0, m_1, m_2} \triangleq \{(s_1^n, s_2^n, U_0^n(m_0), U_1^n(s_1^n, m_1), U_2^n(s_2^n, m_2))$$
$$\in \mathcal{T}_{\epsilon'}^{(n)}\}.$$

Then, taking expectations and using the fact that codewords are independently generated,

$$E[|\mathcal{A}(s_1^n, s_2^n)|] = \sum_{m_0, m_1, m_2} P(\mathcal{A}_{m_0, m_1, m_2})$$
$$= 2^{n(R_0 + R_1 + R_2)} P(\mathcal{A}_{1,1,1}) \qquad (20)$$

Similarly, we have

$$E[|\mathcal{A}(s_1^n, s_2^n)|^2]$$
$$= \sum_{m_0, m_1, m_2} P(\mathcal{A}_{m_0, m_1, m_2})$$
$$+ \sum_{m_0, m_1, m_2} \sum_{m_0' \neq m_0} P(\mathcal{A}_{m_0, m_1, m_2}, \mathcal{A}_{m_0', m_1, m_2})$$
$$+ \sum_{m_0, m_1, m_2} \sum_{m_1' \neq m_1} P(\mathcal{A}_{m_0, m_1, m_2}, \mathcal{A}_{m_0, m_1', m_2})$$
$$+ \sum_{m_0, m_1, m_2} \sum_{m_2' \neq m_2} P(\mathcal{A}_{m_0, m_1, m_2}, \mathcal{A}_{m_0, m_1, m_2'})$$
$$+ \sum_{m_0, m_1, m_2} \sum_{\substack{m_0' \neq m_0 \\ m_1' \neq m_1}} P(\mathcal{A}_{m_0, m_1, m_2}, \mathcal{A}_{m_0', m_1', m_2})$$
$$+ \sum_{m_0, m_1, m_2} \sum_{\substack{m_0' \neq m_0 \\ m_2' \neq m_2}} P(\mathcal{A}_{m_0, m_1, m_2}, \mathcal{A}_{m_0', m_1, m_2'})$$
$$+ \sum_{m_0, m_1, m_2} \sum_{\substack{m_1' \neq m_1 \\ m_2' \neq m_2}} P(\mathcal{A}_{m_0, m_1, m_2}, \mathcal{A}_{m_0, m_1', m_2'})$$
$$+ \sum_{m_0, m_1, m_2} \sum_{\substack{m_0' \neq m_0 \\ m_1' \neq m_1 \\ m_2' \neq m_2}} P(\mathcal{A}_{m_0, m_1, m_2}, \mathcal{A}_{m_0', m_1', m_2'}).$$

Since the codewords are independently generated, we can rewrite the above equality as follows

$$E|\mathcal{A}|^2 \leq 2^{n(R_0 + R_1 + R_2)} P(\mathcal{A}_{1,1,1})$$
$$+ 2^{n(2R_0 + R_1 + R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{2,1,1})$$
$$+ 2^{n(R_0 + 2R_1 + R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{1,2,1})$$
$$+ 2^{n(R_0 + R_1 + 2R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{1,1,2})$$
$$+ 2^{n(2R_0 + 2R_1 + R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{2,2,1})$$
$$+ 2^{n(2R_0 + R_1 + 2R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{2,1,2})$$
$$+ 2^{n(R_0 + 2R_1 + 2R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{1,2,2})$$
$$+ 2^{n(2R_0 + 2R_1 + 2R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{2,2,2}).$$

It is easily seen that $P(\mathcal{A}_{1,1,1}, \mathcal{A}_{2,2,2}) = P(\mathcal{A}_{1,1,1})^2$. It follows that

$$\text{Var}(|\mathcal{A}|) \leq 2^{n(R_0 + R_1 + R_2)} P(\mathcal{A}_{1,1,1})$$
$$+ 2^{n(2R_0 + R_1 + R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{2,1,1})$$
$$+ 2^{n(R_0 + 2R_1 + R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{1,2,1})$$
$$+ 2^{n(R_0 + R_1 + 2R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{1,1,2})$$
$$+ 2^{n(2R_0 + 2R_1 + R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{2,2,1})$$
$$+ 2^{n(2R_0 + R_1 + 2R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{2,1,2})$$
$$+ 2^{n(R_0 + 2R_1 + 2R_2)} P(\mathcal{A}_{1,1,1}, \mathcal{A}_{1,2,2}). \qquad (21)$$

By the joint typicality lemma [8], we have

$$P(\mathcal{A}_{2,1,1}, \mathcal{A}_{1,1,1}) \leq 2^{-n[I(U_0; S_1, S_2, U_1, U_2) + I(U_1; U_0, S_2|S_1)]}$$
$$\cdot 2^{-n[I(U_0; S_1, S_2) + I(U_2; U_0, U_1, S_1|S_2)] - \delta(\epsilon')]},$$
$$P(\mathcal{A}_{1,2,1}, \mathcal{A}_{1,1,1}) \leq 2^{-n[I(U_1; S_2, U_0, U_2|S_1) + I(U_1; U_0, S_2|S_1)]}$$
$$\cdot 2^{-n[I(U_0; S_1, S_2) + I(U_2; U_0, U_1, S_1|S_2)] - \delta(\epsilon')]},$$
$$P(\mathcal{A}_{1,1,2}, \mathcal{A}_{1,1,1}) \leq 2^{-n[I(U_2; S_1, U_0, U_1|S_2) + I(U_1; U_0, S_2|S_1)]}$$
$$\cdot 2^{-n[I(U_0; S_1, S_2) + I(U_2; U_0, U_1, S_1|S_2)] - \delta(\epsilon')]},$$
$$P(\mathcal{A}_{2,2,1}, \mathcal{A}_{1,1,1}) \leq 2^{-n[I(U_0; S_1, S_2, U_1, U_2) + I(U_1; S_2, U_2|S_1)]}$$
$$\cdot 2^{-n[I(U_1; U_0, S_2|S_1) + I(U_0; S_1, S_2)]}$$
$$\cdot 2^{-n[I(U_2; U_0, U_1, S_1|S_2)] - \delta(\epsilon')]},$$
$$P(\mathcal{A}_{2,1,2}, \mathcal{A}_{1,1,1}) \leq 2^{-n[I(U_0; S_1, S_2, U_1, U_2) + I(U_2; S_1, U_1|S_2)]}$$
$$\cdot 2^{-n[I(U_1; U_0, S_2|S_1) + I(U_0; S_1, S_2)]}$$
$$\cdot 2^{-n[I(U_2; U_0, U_1, S_1|S_2)] - \delta(\epsilon')]},$$
$$P(\mathcal{A}_{1,2,2}, \mathcal{A}_{1,1,1}) \leq 2^{-n[I(U_1; U_0, S_2|S_1) + I(U_2; U_0, U_1, S_1|S_2)]}$$
$$\cdot 2^{-n[I(U_1; U_0, S_2|S_1) + I(U_0; S_1, S_2)]}$$
$$\cdot 2^{-n[I(U_2; U_0, U_1, S_1|S_2)] - \delta(\epsilon')]},$$
$$P(\mathcal{A}_{1,1,1}) \geq 2^{-n[I(U_0; S_1, S_2) + I(U_1; U_0, S_2|S_1)}$$
$$\cdot 2^{-n[+I(U_2; U_0, U_1, S_1|S_2)] + \delta(\epsilon')]}.$$

Substituting the above inequalities into (20) and (21), and making use of (19), we obtain that $P(|\mathcal{A}(s_1^n, s_2^n)| = 0) \to 0$ as $n \to \infty$ if conditions (8) are simultaneously satisfied. This

completes the proof of the lemma.

APPENDIX C
ANALYSIS OF THE PROBABILITY OF ERROR

Assume that $(m_0, m_1, m_2) = (M_0, M_1, M_2)$ for the transmitted source sequences (s_1^n, s_2^n). To study the error probability for decoder 1, define the events

$$\mathcal{E}_1 \triangleq \{(S_1^n, U_0^n(M_0), X_1^n(S_1^n, M_0), Y_1^n) \notin \mathcal{T}_\epsilon^{(n)}\},$$
$$\mathcal{E}_2 \triangleq \{(\tilde{s}_1^n, U_0^n(M_0), U_1^n(\tilde{s}_1^n, m_1), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}$$
$$\text{for some } \tilde{s}_1^n \neq S_1^n \text{ and } m_1 \in [1 : 2^{nR_1}]\},$$
$$\mathcal{E}_3 \triangleq \{(\tilde{s}_1^n, U_0^n(m_0), U_1^n(\tilde{s}_1^n, m_1), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}$$
$$\text{for some } \tilde{s}_1^n \neq S_1^n, m_0 \neq M_0, \text{ and } m_1 \in [1 : 2^{nR_1}]\}.$$

The average probability of decoding error at decoder 1 is bounded by

$$\mathsf{P}(\hat{S}_1^n \neq S_1^n) \leq \mathsf{P}(\mathcal{E}_1) + \mathsf{P}(\mathcal{E}_2) + \mathsf{P}(\mathcal{E}_3).$$

By the law of large numbers, $\mathsf{P}(\mathcal{E}_1) \to 0$ as $n \to \infty$. Next, consider the second term. By the union bound, we have

$$\mathsf{P}(\mathcal{E}_2)$$
$$= \sum_{s_1^n} p(s_1^n) P\{(\tilde{s}_1^n, U_0^n(M_0), U_1^n(\tilde{s}_1^n, m_1), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}$$
$$\text{for some } \tilde{s}_1^n \neq S_1^n, \text{ and } m_1 \in [1 : 2^{nR_1}] \mid S_1^n = s_1^n\}$$
$$\leq \sum_{s_1^n} p(s_1^n) \sum_{\tilde{s}_1^n \neq s_1^n} \sum_{m_1=1}^{2^{nR_1}} P\{(\tilde{s}_1^n, U_0^n(M_0), U_1^n(\tilde{s}_1^n, m_1),$$
$$Y_1^n) \in \mathcal{T}_\epsilon^{(n)} \mid S_1^n = s_1^n\}.$$

Conditioned on $S_1^n = s_1^n$, for all $\tilde{s}_1^n \neq s_1^n$ and for all $m_1 \in [1 : 2^{nR_1}]$, we have that $(U_0^n(M_0), U_1^n(\tilde{s}_1^n, m_1), Y_1^n) \sim p(u_1^n(\tilde{s}_1^n, m_1)|\tilde{s}_1^n) \, p(u_0^n(M_0), y_1^n|s_1^n)$. Thus,

$$\mathsf{P}(\mathcal{E}_2)$$
$$\leq \sum_{s_1^n} p(s_1^n) \sum_{\tilde{s}_1^n \neq s_1^n, (\tilde{s}_1^n, u_0^n, u_1^n, y_1^n) \in \mathcal{T}_\epsilon^{(n)}} \sum_{m_1=1}^{2^{nR_1}} p(u_0^n(M_0), y_1^n|s_1^n)$$
$$\cdot p(u_1^n(\tilde{s}_1^n, m_1)|\tilde{s}_1^n)$$
$$= \sum_{(\tilde{s}_1^n, u_0^n, u_1^n, y_1^n) \in \mathcal{T}_\epsilon^{(n)}} \sum_{s_1^n \neq \tilde{s}_1^n} \sum_{m_1=1}^{2^{nR_1}} p(u_0^n(M_0), y_1^n|s_1^n)$$
$$\cdot p(u_1^n(\tilde{s}_1^n, m_1)|\tilde{s}_1^n) p(s_1^n)$$
$$\leq \sum_{(\tilde{s}_1^n, u_0^n, u_1^n, y_1^n) \in \mathcal{T}_\epsilon^{(n)}} \sum_{s_1^n} \sum_{m_1=1}^{2^{nR_1}} p(u_0^n(M_0), y_1^n|s_1^n)$$
$$\cdot p(u_1^n(\tilde{s}_1^n, m_1)|\tilde{s}_1^n) p(s_1^n)$$
$$= \sum_{(\tilde{s}_1^n, u_0^n, u_1^n, y_1^n) \in \mathcal{T}_\epsilon^{(n)}} \sum_{m_1=1}^{2^{nR_1}} p(u_0^n(M_0), y_1^n) p(u_1^n(\tilde{s}_1^n, m_1)|\tilde{s}_1^n)$$
$$\leq \sum_{(\tilde{s}_1^n, u_0^n, u_1^n, y_1^n) \in \mathcal{T}_\epsilon^{(n)}} 2^{nR_1} 2^{-n(H(U_0, Y_1) + H(U_1|S_1) - 2\delta(\epsilon))}$$
$$\leq 2^{-n(H(S_1, U_0, U_1, Y_1) + \delta(\epsilon)) + nR_1 - n(H(U_0, Y_1) + H(U_1|S_1) - 2\delta(\epsilon))}.$$

Collecting the entropy terms at the exponent, we have

$$H(S_1, U_0, U_1, Y_1) - H(U_0, Y_1) - H(U_1|S_1)$$
$$= H(S_1) + H(U_1|S_1) + H(U_0, Y_1|U_1, S_1)$$
$$\quad - H(U_0, Y_1) - H(U_1|S_1)$$
$$= H(S_1) - I(U_1, S_1; U_0, Y_1).$$

Thus $\mathsf{P}(\mathcal{E}_2) \to 0$ as $n \to \infty$ if

$$H(S_1) + R_1 < I(U_1, S_1; U_0, Y_1) - 2\delta(\epsilon). \tag{22}$$

Finally, consider the third term. By the union bound, we have

$$\mathsf{P}(\mathcal{E}_3)$$
$$= \sum_{s_1^n} p(s_1^n) P\{(\tilde{s}_1^n, U_0^n(m_0), U_1^n(\tilde{s}_1^n, m_1), Y_1^n) \in \mathcal{T}_\epsilon^{(n)}$$
$$\text{for some } \tilde{s}_1^n \neq S_1^n, m_0 \neq M_0, \text{ and } m_1|S_1^n = s_1^n\}$$
$$\leq \sum_{s_1^n} p(s_1^n) \sum_{\tilde{s}_1^n \neq s_1^n} \sum_{m_0=1}^{2^{nR_0}} \sum_{m_1=1}^{2^{nR_1}} P\{(\tilde{s}_1^n, U_0^n(m_0), U_1^n(\tilde{s}_1^n, m_1),$$
$$Y_1^n) \in \mathcal{T}_\epsilon^{(n)}|S_1^n = s_1^n\}.$$

Conditioned on $S_1^n = s_1^n$, for all $\tilde{s}_1^n \neq s_1^n$, $m_0 \neq M_0$, and for all m_1, we have that $(U_0^n(m_0), U_1^n(\tilde{s}_1^n, m_1), Y_1^n) \sim p(u_0^n(m_0))p(u_1^n(\tilde{s}_1^n, m_1)|\tilde{s}_1^n)p(y_1^n|s_1^n)$. Thus,

$$\mathsf{P}(\mathcal{E}_3)$$
$$\leq \sum_{s_1^n} p(s_1^n) \sum_{\tilde{s}_1^n \neq s_1^n, (\tilde{s}_1^n, u_0^n, u_1^n, y_1^n) \in \mathcal{T}_\epsilon^{(n)}} \sum_{m_0=1}^{2^{nR_0}} \sum_{m_1=1}^{2^{nR_1}} p(u_0^n(m_0))$$
$$\cdot p(u_1^n(\tilde{s}_1^n, m_1)|\tilde{s}_1^n)p(y_1^n|s_1^n)$$
$$= \sum_{(\tilde{s}_1^n, u_0^n, u_1^n, y_1^n) \in \mathcal{T}_\epsilon^{(n)}} \sum_{s_1^n \neq \tilde{s}_1^n} \sum_{m_0=1}^{2^{nR_0}} \sum_{m_1=1}^{2^{nR_1}} p(u_0^n(m_0))$$
$$\cdot p(u_1^n(\tilde{s}_1^n, m_1)|\tilde{s}_1^n)p(y_1^n|s_1^n)p(s_1^n)$$
$$\leq \sum_{(\tilde{s}_1^n, u_0^n, u_1^n, y_1^n) \in \mathcal{T}_\epsilon^{(n)}} \sum_{s_1^n} \sum_{m_0=1}^{2^{nR_0}} \sum_{m_1=1}^{2^{nR_1}} p(u_0^n(m_0))$$
$$\cdot p(u_1^n(\tilde{s}_1^n, m_1)|\tilde{s}_1^n)p(y_1^n|s_1^n)p(s_1^n)$$
$$= \sum_{(\tilde{s}_1^n, u_0^n, u_1^n, y_1^n) \in \mathcal{T}_\epsilon^{(n)}} \sum_{m_0=1}^{2^{nR_0}} \sum_{m_1=1}^{2^{nR_1}} p(u_0^n(m_0))$$
$$\cdot p(u_1^n(\tilde{s}_1^n, m_1)|\tilde{s}_1^n)p(y_1^n)$$
$$\leq \sum_{(\tilde{s}_1^n, u_0^n, u_1^n, y_1^n) \in \mathcal{T}_\epsilon^{(n)}} 2^{n[R_0 + R_1 - (H(U_0) + H(U_1|S_1) + H(Y_1) - 3\delta(\epsilon))]}$$
$$\leq 2^{-n(H(S_1, U_0, U_1, Y_1) + \delta(\epsilon))} 2^{n(R_0 + R_1)}$$
$$\cdot 2^{-n(H(U_0) + H(U_1|S_1) + H(Y_1) - 3\delta(\epsilon))}.$$

Collecting the entropy terms at the exponent, we have

$$H(S_1, U_0, U_1, Y_1) - H(U_0) - H(U_1|S_1) - H(Y_1)$$
$$= H(S_1) + H(U_1|S_1) + H(U_0|U_1, S_1) + H(Y_1|S_1, U_0, U_1)$$
$$\quad - H(U_0) - H(U_1|S_1) - H(Y_1)$$
$$= H(S_1) - I(U_0, U_1, S_1; Y_1) - I(U_0; U_1, S_1).$$

Thus $P(\mathcal{E}_3) \to 0$ as $n \to \infty$ if

$$H(S_1) + R_0 + R_1 < I(U_0, U_1, S_1; Y_1)+$$
$$+ I(U_0; U_1, S_1) - 3\delta(\epsilon). \quad (23)$$

In summary, the probability of error for decoder 1 can be made arbitrarily small by letting $n \to \infty$ if (22) and (23) hold. Finally, the study of the error probability for decoder 2 follows from a similar argument.

APPENDIX D
AN ALTERNATIVE CODING SCHEME USING SUPERPOSITION CODING

We describe here the construction of coding scheme based on superposition coding which yields the same sufficient condition as the one stated in Theorem 2.

Random codebook generation: Let $\epsilon' > 0$. Fix a joint distribution $p(u_0, u_1, u_2|s_1, s_2)$ and, without loss of generality, let $p(x|u_0, u_1, u_2, s_1, s_2)$ be a chosen deterministic function $x(s_1, s_2, u_0, u_1, u_2)$. Compute $p(u_0)$, $p(u_1|u_0, s_1)$ and $p(u_2|u_0, s_2)$ for the given source distribution $p(s_1, s_2)$. Randomly and independently generate 2^{nR_0} sequences $u_0^n(m_0)$, $m_0 \in [1 : 2^{nR_0}]$, each according to $\prod_{i=1}^{n} p_{U_0}(u_{0i})$. For each source sequence s_1^n and $u_0^n(m_0)$, randomly and independently generate 2^{nR_1} sequences $u_1^n(s_1^n, m_0, m_1)$, $m_1 \in [1 : 2^{nR_1}]$, each according to $\prod_{i=1}^{n} p_{U_1|S_1,U_0}(u_{1,i}|s_{1i}, u_{0,i}(m_0))$. Similarly, for each source sequence s_2^n and $u_0^n(m_0)$, randomly and independently generate 2^{nR_2} sequences $u_2^n(s_2^n, m_0, m_2)$, $m_2 \in [1 : 2^{nR_2}]$, each according to $\prod_{i=1}^{n} p_{U_2|S_2,U_0}(u_{2,i}|s_{2,i}, u_{0,i}(m_0))$. The rates (R_0, R_1, R_2) are chosen so that the ensemble of generated sequences "cover" the set $\mathcal{T}_{\epsilon'}^{(n)}(U_0, U_1, U_2|s_1^n, s_2^n)$ for all $s_1^n, s_2^n \in \mathcal{T}_{\epsilon'}^{(n)}(S_1, S_2)$. Define the event

$$\mathcal{A} = \{(S_1^n, S_2^n, U_0^n(m_0), U_1^n(S_1^n, m_0, m_1),$$
$$U_2^n(S_2^n, m_0, m_2)) \notin \mathcal{T}_{\epsilon'}^{(n)} \text{ for all } m_0, m_1, m_2\}$$

It can be shown using techniques similar to those used in the proof of the Lemma 5 that $P(\mathcal{A}) \to 0$ as $n \to \infty$ if

$$R_0 > I(U_0; S_1, S_2) + \delta(\epsilon'),$$
$$R_0 + R_1 > I(U_0; S_1, S_2) + I(S_2; U_1|S_1, U_0) + \delta(\epsilon'),$$
$$R_0 + R_2 > I(U_0; S_1, S_2) + I(S_1; U_2|S_2, U_0) + \delta(\epsilon'),$$
$$R_0 + R_1 + R_2 > I(U_0; S_1, S_2) + I(S_2; U_1|S_1, U_0)$$
$$+ I(S_1, U_1; U_2|S_2, U_0) + \delta(\epsilon') \quad (24)$$

where $\delta(\epsilon') \to 0$ as $\epsilon' \to 0$.

Encoding: For each source sequence (s_1^n, s_2^n), choose a triple $(m_0, m_1, m_2) \in [1 : 2^{nR_0}] \times [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ such that $(s_1^n, s_2^n, u_0^n(m_0), u_1^n(s_1^n, m_0, m_1), u_2^n(s_2^n, m_0, m_2)) \in \mathcal{T}_{\epsilon'}^{(n)}(S_1, S_2, U_0, U_1, U_2)$. If there is no such triple, choose $(m_0, m_1, m_2) = (1, 1, 1)$. Then at time $i \in [1 : n]$, the encoder transmits $x_i = x(s_{1i}, s_{2i}, u_{0i}(m_0), u_{1i}(s_1^n, m_0, m_1), u_{2i}(s_2^n, m_0, m_2))$.

Decoding: Let $\epsilon > 0$. Decoder 1 declares $\hat{s}_1^n$ to be the estimate of the source s_1^n if it is the unique sequence such that $(\hat{s}_1^n, u_0^n(m_0), u_1^n(\hat{s}_1^n, m_0, m_1), y_1^n) \in \mathcal{T}_{\epsilon}^{(n)}(S_1, U_0, U_1, Y_1)$ for some $(m_0, m_1) \in [1 : 2^{nR_0}] \times [1 : 2^{nR_1}]$. Similarly, decoder 2 declares $\hat{s}_2^n$ to be the estimate of

the source u_2^n if it is the unique sequence such that $(\hat{s}_2^n, u_0^n(m_0), u_2^n(\hat{s}_2^n, m_0, m_2), y_2^n) \in \mathcal{T}_{\epsilon}^{(n)}(S_2, U_0, U_2, Y_2)$ for some $(m_0, m_2) \in [1 : 2^{nR_0}] \times [1 : 2^{nR_2}]$.

Error events: Assume that $(m_0, m_1, m_2) = (M_0, M_1, M_2)$ is selected for the source sequence (s_1^n, s_2^n). The error event for decoder 1 can be divided into two parts:

1) There exists a sequence $\hat{s}_1^n \neq S_1^n$ such that $(\hat{s}_1^n, u_0^n(M_0), u_1^n(\hat{s}_1^n, M_0, m_1), y_1^n) \in \mathcal{T}_{\epsilon}^{(n)}(S_1, U_0, U_1, Y_1)$ for some $m_1 \in [1 : 2^{nR_1}]$. The probability of this event vanishes as $n \to \infty$ if the following inequality is satisfied

$$H(S_1) + R_1 < I(U_1, S_1; Y_1|U_0) + I(S_1; U_0) - \delta(\epsilon). \quad (25)$$

2) There exists a sequence $\hat{s}_1^n \neq S_1^n$ and an $m_0 \neq M_0$ such that $(\hat{s}_1^n, u_0^n(m_0), u_1^n(\hat{s}_1^n, m_0, m_1), y_1^n) \in \mathcal{T}_{\epsilon}^{(n)}(S_1, U_0, U_1, Y_1)$ for some $m_1 \in [1 : 2^{nR_1}]$. The probability of this event can be made arbitrarily small as $n \to \infty$ if

$$H(S_1) + R_0 + R_1 < I(U_0, U_1, S_1; Y_1)+$$
$$+ I(U_0; S_1) - \delta(\epsilon). \quad (26)$$

Similarly, the probability of error for decoder 2 can be made arbitrarily small as $n \to \infty$ if

$$H(S_2) + R_2 < I(U_2, S_2; Y_2|U_0) + I(S_2; U_0) - \delta(\epsilon),$$
$$H(S_2) + R_0 + R_2 < I(U_0, U_2, S_2; Y_2) + I(U_0; S_2) - \delta(\epsilon). \quad (27)$$

The rest of the proof follows by letting $\epsilon', \epsilon \to 0$, then $n \to \infty$, and finally by eliminating (R_0, R_1, R_2) from (24), (25), (26), and (27) using the Fourier–Motzkin elimination algorithm.

REFERENCES

[1] T. Christof, A. Löbel, PORTA – Polyhedron Representation Transformation Algorithm, http://www.zib.de/Optimization/Software/Porta/
[2] T. M. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 2–14, Jan. 1972.
[3] T. Cover, "An achievable rate region for the broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-21, no. 4, pp. 399–404, Jul. 1975.
[4] T. M. Cover, A. El Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 6, pp. 648–657, Nov. 1980.
[5] I. Csiszár and J. Körner, *Information Theory: coding theorems for discrete memoryless systems*. Budapest, Hungary: Akadmiai Kiad, 1981.
[6] A. El Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 166-169, Mar. 1979.
[7] A. El Gamal and E. C. van der Meulen, "A proof of Marton's coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 120–122, Jan. 1981.
[8] A. El Gamal and Y. H. Kim, *Lecture notes on network information theory*.
[9] W. Kang and G. Kramer, "Broadcast channel with degraded source random variables and receiver side information," *in IEEE Int. Symp. Inform. Theory,*, Toronto, Canada, pp.1711-1715, July 2008.
[10] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 1, pp. 60-64, Jan. 1977.
[11] G. Kramer and C. Nair, "Comments on 'Broadcast channels with arbitrarily correlated sources'," http://arxiv.org/abs/0901.0220.
[12] P. Gáks and J. Körner, "Common information is much less than mutual information," *Problems Contr. Inform. Theory*, vol. 2, pp. 149–162, 1973.
[13] S. I. Gelfand and M. S. Pinsker, "Capacity of a broadcast channel with one deterministic component," *Probl. Inform. Transm.*, vol. 16, no. 1, pp. 17–25, Jan.-Mar. 1980.

[14] A. A. Gohari and V. Anantharam, "An outer bound to the admissible source region of broadcast channels with arbitrarily correlated sources and channel variations," *Proceedings of the 46th Annual Allerton Conference on Communications, Control and Computing*, September 2008.

[15] R. M. Gray and A. D. Wyner, "Source coding for a simple network," *Bell System Tech. J.*, vol. IT-53, pp. 1681–1721, 1974.

[16] T. S. Han and M. H. M. Costa, "Broadcast channels with arbitrarily correlated sources," *IEEE Trans. Inform. Theory*, vol. IT-33, no. 5, pp. 641–650, Sep. 1987.

[17] T. S. Han and K. Kobayashi,"A new achievable rate region for the interference channel," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 1, pp. 49–60, Jan 1981.

[18] G. Kramer, Y. Liang and S. Shamai (Shitz), "Outer bounds on the admissible source region for broadcast channels with dependent sources", *Inform. Theory and Applications Workshop*, 2009.

[19] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *IEEE Trans. Inform. Theory*, vol. 53, no. 10, pp. 3517–3535, Oct. 2007.

[20] Y. Liang, G. Kramer, and S. Shamai (Shitz), "Capacity outer bounds for broadcast channels," *IEEE Inform. Theory Workshop*, Porto, Portugal, pp. 2-4, May 5-9, 2008.

[21] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 3, pp. 306–311, May 1979.

[22] C. Nair, "An outer bound for 2-receiver discrete memoryless broadcast channels," http://chandra.ie.cuhk.edu.hk/pub/papers/outerbound.pdf

[23] C. Nair and A. A. El Gamal, "An outer bound to the capacity region of the broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-53, no. 1, pp. 350–355, Jan. 2007.

[24] E. C. van der Meulen, "Random coding theorems for the general discrete memoryless broadcast channel," *IEEE Trans. Inform. Theory*, vol. IT-21, no. 2, pp. 180–190, Mar. 1975.

[25] R. W. Yeung and Y.-O. Yan, ITIP - Information theoretic inequality prover, http://www.ie.cuhk.edu.hk/~ITIP.

[26] H. S. Witsenhausen, "On sequences of pairs of dependent random variables," *SIAM J. Appl. Math.*, vol. 28, pp. 100–113, Jan. 1975.

[27] G. M. Ziegler, *Lectures on Polytopes*, vol. 152 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1995.